

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2006, 09:27:05 ; Search time 27 Seconds  
(without alignments)  
1886.230 Million cell updates/sec

Title: US-10-734-782-4

Perfect score: 3097

Sequence: 1 MADKGMTYSPDVRNTLVRR.....DMTPRIFRDEDEEDIDI 616

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgm2\_6/ptodata/1/iaa/5\_COMB.pep.\*
- 2: /cgm2\_6/ptodata/1/iaa/6\_COMB.pep.\*
- 3: /cgm2\_6/ptodata/1/iaa/H\_COMB.pep.\*
- 4: /cgm2\_6/ptodata/1/iaa/FCRUS\_COMB.pep.\*
- 5: /cgm2\_6/ptodata/1/iaa/RG\_COMB.pep.\*
- 6: /cgm2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3097	100.0	616	2	US-10-049-086-4
2	3015	97.4	616	2	US-09-690-185A-2
3	128	4.1	1095	2	US-09-206-942-45
4	128	4.1	1101	2	US-09-206-942-43
5	118	3.8	564	2	US-09-107-532A-6970
6	116.5	3.8	2042	2	US-09-077-098A-6
7	116.5	3.8	2042	2	US-10-192-584-6
8	116	3.7	1166	2	US-09-200-6508-7
9	116	3.7	1781	1	US-08-477-451-11
10	115	3.7	1095	2	US-09-206-942-69
11	115	3.7	1095	2	US-10-193-764-65
12	115	3.7	1536	1	US-08-038-682-2
13	115	3.7	1536	1	US-08-302-832-2
14	115	3.7	1536	1	US-08-530-198-2
15	115	3.7	1536	1	US-08-469-880-2
16	115	3.7	1536	1	US-08-728-470-2
17	115	3.7	1536	1	US-08-617-697-2
18	115	3.7	1536	2	US-08-719-641-2
19	115	3.7	1536	2	US-10-193-764-63
20	115	3.7	1536	2	US-09-206-942-67
21	114	3.7	2048	2	US-09-268-347-48
22	114	3.7	3241	2	US-09-841-786-1
23	113	3.6	504	2	US-09-949-016-7935
24	112.5	3.6	1572	2	US-09-710-279-2306
25	112.5	3.6	4536	2	US-09-180-422B-27
26	112.5	3.6	4536	2	US-09-079-030-1
27	112.5	3.6	4563	2	US-09-108-006C-1

28	112.5	3.6	4563	2	US-09-538-092-842	Sequence 842, Appl
29	111.5	3.6	1312	2	US-09-345-882-29	Sequence 29, Appl
30	111.5	3.6	1312	2	US-10-071-179-29	Sequence 29, Appl
31	111	3.6	789	2	US-09-134-000C-4939	Sequence 4939, Ap
32	111	3.6	1833	2	US-08-621-944A-4	Sequence 4, Appl
33	111	3.6	1833	2	US-08-945-567D-4	Sequence 4, Appl
34	111	3.6	1992	2	US-08-621-944A-3	Sequence 3, Appl
35	111	3.6	1992	2	US-08-945-567D-3	Sequence 3, Appl
36	110.5	3.6	1129	2	US-09-252-991A-29927	Sequence 29927, A
37	110	3.6	2123	2	US-08-968-685A-10	Sequence 10, Appl
38	109.5	3.5	1477	1	US-08-038-682-4	Sequence 4, Appl
39	109.5	3.5	1477	1	US-08-302-832-4	Sequence 4, Appl
40	109.5	3.5	1477	1	US-08-530-198-4	Sequence 4, Appl
41	109.5	3.5	1477	1	US-08-469-880-4	Sequence 4, Appl
42	109.5	3.5	1477	1	US-08-728-470-4	Sequence 4, Appl
43	109.5	3.5	1477	1	US-08-617-697-4	Sequence 4, Appl
44	109.5	3.5	1477	2	US-08-719-641-4	Sequence 4, Appl
45	109	3.5	1004	2	US-09-206-942-57	Sequence 57, Appl

## ALIGNMENTS

RESULT 1  
US-10-049-086-4  
; Sequence 4, Application US/10049086  
; Patent No. 6919083  
; GENERAL INFORMATION:  
; APPLICANT: Aqua Health (Europe) Limited  
; APPLICANT: Griffiths, Steven  
; APPLICANT: Ritchie, Rachael  
; TITLE OF INVENTION: Sequence  
; FILE REFERENCE: P24268-/GST/RMC  
; CURRENT APPLICATION NUMBER: US/10/049,086  
; PRIOR FILING DATE: 2002-02-06  
; PRIOR APPLICATION NUMBER: GB9918588.6  
; PRIOR FILING DATE: 1999-07-08  
; PRIOR APPLICATION NUMBER: GB0005848.7  
; PRIOR FILING DATE: 2000-03-11  
; PRIOR APPLICATION NUMBER: GB0006674.6  
; PRIOR FILING DATE: 2000-03-21  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 4  
; LENGTH: 616  
; TYPE: EPT  
; ORGANISM: Salmon Anaemia Virus  
US-10-049-086-4

Query Match	100.0%;	Score 3097;	DB 2;	Length 616;	
Best Local Similarity	100.0%;	Pred. No. 3.7e-309;	Mismatches 0;	Indels 0;	Gaps 0;
Matches 616;	Conservative 0;				
Qy	1	MADKGMTYSPDVRNTLVRRSTATKSGIKISYREDRTGTSLLQKAFAGTDEDFWVLPDQD	60		
Db	1	MADKGMTYSPDVRNTLVRRSTATKSGIKISYREDRTGTSLLQKAFAGTDEDFWVLPDQD	60		
Qy	61	VYVDKIKRFLBEEKMDMSTRVSGAVAAAIERSVDFPNFSEKAAANTEMAGVDDERAGG	120		
Db	61	VYVDKIKRFLBEEKMDMSTRVSGAVAAAIERSVDFPNFSEKAAANTEMAGVDDERAGG	120		
Qy	121	SGLYDNRKNGKGVSNMAYNLSLFTGMVFPALTTFPSATLSGEMSIWONGOAIIRILALA	180		
Db	121	SGLYDNRKNGKGVSNMAYNLSLFTGMVFPALTTFPSATLSGEMSIWONGOAIIRILALA	180		
Qy	181	DEDKRQTRTGQQRVDMADVTKLVNVTANGVKQVEVNLNDLKAAPRQSRPERSDYRKQ	240		
Db	181	DEDKRQTRTGQQRVDMADVTKLVNVTANGVKQVEVNLNDLKAAPRQSRPERSDYRKQ	240		
Qy	241	GSKATESISNQCMALIMKSVLSADQIFAPGVKWRNGFNASTYTTLAEGANIPSKYLH	300		
Db	241	GSKATESISNQCMALIMKSVLSADQIFAPGVKWRNGFNASTYTTLAEGANIPSKYLH	300		

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2006, 09:27:06 ; Search time 65 Seconds  
(without alignments)  
3959.736 Million cell updates/sec

Title: US-10-734-782-4  
Perfect score: 3097  
Sequence: 1 MADKGMTYSFVDRDNTLVVR.....DMPRIEFDEDEEDIDI 616

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA Main:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pap:  
2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pap:  
3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pap:  
4: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pap:  
5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pap:  
6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3097	100.0	616	4	US-10-734-782-4
2	3097	100.0	616	6	US-11-083-801-4
3	3083	99.5	616	4	US-10-472-990-6
4	128	4.1	1556	4	US-10-369-493-1796
5	126	4.1	957	4	US-10-607-631-12
6	121.5	3.9	954	5	US-10-450-763-52631
7	121.5	3.9	1377	4	US-10-282-122A-47115
8	121	3.9	2402	4	US-10-661-809-20
9	119	3.8	735	4	US-10-767-701-46665
10	119	3.8	872	4	US-10-032-585-7421
11	118.5	3.8	1483	4	US-10-282-122A-51483
12	117	3.8	1226	5	US-10-471-934-2
13	117	3.8	1226	5	US-10-471-934-16
14	116.5	3.8	1166	4	US-10-282-122A-51888
15	116.5	3.8	2042	4	US-10-192-584-6
16	116	3.7	839	4	US-10-369-493-13153
17	116	3.7	1166	5	US-10-744-672-7
18	116	3.7	1166	5	US-10-744-616-7
19	116	3.7	1166	5	US-10-470-048B-153
20	116	3.7	2516	4	US-10-437-963-162525
21	115.5	3.7	1191	4	US-10-282-122A-52048
22	115	3.7	1095	4	US-10-193-764-65
23	115	3.7	1536	4	US-10-092-880-2
24	115	3.7	1536	4	US-10-193-764-63
25	115	3.7	1536	5	US-10-681-171-2
26	114.5	3.7	1086	4	US-10-282-122A-74458
27	114.5	3.7	4560	5	US-10-398-200-2

28	114	3.7	3241	3	US-09-841-786-1	Sequence 1, Appli
29	114	3.7	3241	4	US-10-647-057-1	Sequence 1, Appli
30	113.5	3.7	368	4	US-10-156-761-12636	Sequence 12636, A
31	113.5	3.7	932	5	US-10-450-763-40981	Sequence 40981, A
32	113.5	3.7	1312	4	US-10-221-625-95	Sequence 95, Appl
33	113.5	3.7	2052	4	US-10-282-122A-51602	Sequence 51602, A
34	113	3.6	906	4	US-10-282-122A-57631	Sequence 57631, A
35	112.5	3.6	1706	3	US-09-864-761-46862	Sequence 46862, A
36	112.5	3.6	2139	5	US-10-480-456-1	Sequence 1, Appli
37	112.5	3.6	3000	4	US-10-741-601-431	Sequence 431, App
38	112.5	3.6	3000	5	US-10-741-600-1286	Sequence 1286, Ap
39	112.5	3.6	4536	5	US-10-656-053B-1	Sequence 1, Appli
40	112.5	3.6	4563	3	US-09-870-759-128	Sequence 128, App
41	112.5	3.6	4563	3	US-09-802-640-32	Sequence 32, Appl
42	112.5	3.6	4563	3	US-09-751-708A-128	Sequence 128, App
43	112.5	3.6	4563	4	US-10-403-902A-32	Sequence 32, Appl
44	112.5	3.6	4563	4	US-10-741-601-432	Sequence 432, App
45	112.5	3.6	4563	4	US-10-741-601-433	Sequence 433, App

## ALIGNMENTS

RESULT 1  
US-10-734-782-4  
; Sequence 4, Application US/10734782  
; Publication No. US20040146860A1  
; GENERAL INFORMATION:  
; APPLICANT: Griffiths, Steven  
; TITLE OF INVENTION: Nucleic Acid And Amino Acid Sequences Of  
; TITLE OF INVENTION: Infectious Salmon Anaemia Virus And Their Uses As Vaccines  
; FILE REFERENCES: H-323188  
; CURRENT APPLICATION NUMBER: US/10/734,782  
; PRIOR FILING DATE: 2003-12-12  
; PRIOR APPLICATION NUMBER: US 10/049,086  
; PRIOR FILING DATE: 2002-02-06  
; PRIOR APPLICATION NUMBER: PCT/GB00/02976  
; PRIOR FILING DATE: 2000-08-07  
; PRIOR APPLICATION NUMBER: GB 0006674.6  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: GB 0005948.7  
; PRIOR FILING DATE: 2000-03-11  
; PRIOR APPLICATION NUMBER: GB 9918588.6  
; PRIOR FILING DATE: 1999-08-07  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 616  
; TYPE: PRT  
; ORGANISM: Infectious Salmon Anaemia Virus  
US-10-734-782-4

Query Match	100.0%;	Score 3097;	DB 4;	Length 616;
Best Local Similarity	100.0%;	Pred. No. 6.8e-270;		
Matches	616;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	1	MADKGMTYSFVDRDNTLVVRSTATSGIKISYREDRGTSLLQKAPAGTEDEFWVLDQD	60	
DB	1	MADKGMTYSFVDRDNTLVVRSTATSGIKISYREDRGTSLLQKAPAGTEDEFWVLDQD	60	
QY	61	VVVDKKIRKFLBEEKKDKMTSVSGAVAAAIERSVEFDNFSKEAAANIEMAGVDDREAGG	120	
DB	61	VVVDKKIRKFLBEEKKDKMTSVSGAVAAAIERSVEFDNFSKEAAANIEMAGVDDREAGG	120	
QY	121	SLGVNRRKNGVSNWYNVLSLFGVFPALTTFFSAILSEGMSTWQGOAIRILALA	180	
DB	121	SLGVNRRKNGVSNWYNVLSLFGVFPALTTFFSAILSEGMSTWQGOAIRILALA	180	
QY	181	DEGKRQRTTGQQRVDMADVTKLNVVVTANGKVKQVEVNLNDLKAAFPQSRPKESDYRKQG	240	
DB	181	DEGKRQRTTGQQRVDMADVTKLNVVVTANGKVKQVEVNLNDLKAAFPQSRPKESDYRKQG	240	

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2006, 09:27:05 ; Search time 80 Seconds  
(without alignments)  
3383.217 Million cell updates/sec

Title: US-10-734-782-4

Perfect score: 3097

Sequence: 1 MADKGMTYSFVDRNTLVVR.....DMTPRIEFDEDEEDIDI 616

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: A Geneseq 21.\*

- 1: geneseqp1980s.\*
- 2: geneseqp1990s.\*
- 3: geneseqp2000s.\*
- 4: geneseqp2001s.\*
- 5: geneseqp2002s.\*
- 6: geneseqp2003as.\*
- 7: geneseqp2003bs.\*
- 8: geneseqp2004s.\*
- 9: geneseqp2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3097	100.0	616	4	AAB71827	Aab71827 Infectiou
2	3083	99.5	616	6	AAB32454	Aab32454 ISAV NP P
3	3034	98.0	616	4	AAB47509	Aab47509 ISA anti
4	3020	97.5	616	4	AAB07613	Aab07613 ISAV stru
5	744	24.0	159	4	AAB47510	Aab47510 ISA anti
6	721	23.3	148	4	AAB47511	Aab47511 ISA anti
7	128.5	4.1	6077	6	ABP95996	Abp95996 White spo
8	128	4.1	1095	3	AAB01835	Aab01835 Haemophil
9	128	4.1	1101	3	AAB01834	Aab01834 Haemophil
10	128	4.1	1556	8	ADN19143	Adn19143 Bacterial
11	126	4.1	957	8	ADT28619	Adt28619 Mycoplasma
12	123	4.0	1314	8	ADR88903	Adr88903 Anopheles
13	121.5	3.9	954	4	ABG22272	Abg22272 Novel hum
14	121.5	3.9	1377	6	ABU19191	Abu19191 Protein e
15	121	3.9	2402	8	AD084848	Ad084848 S epiderm
16	120.5	3.9	959	8	ADM67692	Adm67692 Tobacco N
17	120.5	3.9	959	8	ADT04785	Adt04785 Common to
18	119	3.8	872	5	ABP73584	Abp73584 Candida a
19	119	3.8	1371	3	AAG54221	Aag54221 Arabidops
20	119	3.8	1377	3	AAG54220	Aag54220 Arabidops
21	119	3.8	1443	3	AAG54219	Aag54219 Arabidops
22	118.5	3.8	1483	6	ABU23559	Abu23559 Protein e
23	118	3.8	564	7	ADC97343	Adc97343 E. faeciu
24	118	3.8	907	4	AAM40087	Aam40087 Human pol

25	117	3.8	1226	5	ABP54852	Abp54852 Human ret
26	117	3.8	1536	2	AAR41725	Aar41725 High mole
27	116.5	3.8	1086	9	AEB91670	Aeb91670 Microbial
28	116.5	3.8	1166	6	ABU23964	Abu23964 Protein e
29	116.5	3.8	1600	9	AEB91199	Aeb91199 Outer mem
30	116.5	3.8	1719	9	AEB91201	Aeb91201 Outer mem
31	116.5	3.8	2042	2	AAS56319	Aas56319 Haemophil
32	116	3.7	839	8	ADS24120	Ads24120 Bacterial
33	116	3.7	1166	6	AAY08643	Aay08643 S. aureus
34	116	3.7	1166	6	ABJ18982	Abj18982 Pathogen
35	116	3.7	1226	5	ABP54847	Abp54847 Human ret
36	116	3.7	1606	8	ADM99231	Adm99231 Environme
37	115.5	3.7	1191	6	ABU24124	Abu24124 Protein e
38	115.5	3.7	1309	9	ADW76023	Adw76023 Ralstonia
39	115	3.7	748	3	AAB42613	Aab42613 Human ORF
40	115	3.7	1095	3	AAB01847	Aab01847 Haemophil
41	115	3.7	1536	2	AAR41723	Aar41723 High mole
42	115	3.7	1536	2	AAR63505	Aar63505 Haemophil
43	115	3.7	1536	3	AAB01846	Aab01846 Haemophil
44	114.5	3.7	594	3	AAG47035	Aag47035 Arabidops
45	114.5	3.7	615	3	AAG47034	Aag47034 Arabidops

## ALIGNMENTS

RESULT 1	
AAB71827	
ID	AAB71827 standard; protein; 616 AA.
XX	
AC	AAB71827;
XX	
DT	11-SEP-2003 (revised)
DT	02-MAY-2001 (first entry)
XX	
DE	Infectious salmon anaemia virus ISALmta protein.
XX	
KW	Infectious salmon anaemia virus; ISAV; ISALmta; antiviral; vaccine.
XX	
OS	Infectious salmon anemia virus.
XX	
FN	WO200110469-A2.
XX	
PD	15-FEB-2001.
XX	
PF	07-AUG-2000; 2000WO-GB002976.
XX	
PR	07-AUG-1999; 99GB-00018588.
PR	11-MAR-2000; 2000GB-00005848.
PR	21-MAR-2000; 2000GB-00006674.
XX	
XX	(AQUA-) AQUA HEALTH EURO LTD.
XX	
PI	Griffiths S, Ritchie RJ;
XX	
DR	WPI; 2001-191498/19.
DR	N-PSDB; AAP62721.
XX	
PT	Composition containing nucleic acid or protein derived from infectious salmon anemia virus, useful as vaccine for protecting fish against infection.
XX	
PS	Claim 3; Fig 4; 45pp; English.
XX	
CC	The present invention provides a composition containing at least one nucleic acid and/or amino acid sequence, or their synthetic analogues or homologues, derived from the sequences of infectious salmon anaemia virus (ISAV). The sequences do not cause salmon anaemia and can be used as, or to prepare, a vaccine against ISAV. The composition may also be used to design more accurate diagnostic tests, e.g. for epidemiological studies on the dissemination of different viral strains. The present amino acid sequence may be used in the composition. (Updated on 11-SEP-2003 to standardise OS field)

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2006, 09:27:05 ; Search time 20 Seconds  
(without alignments)  
2963.476 Million cell updates/sec

Title: US-10-734-782-4  
Perfect score: 3097  
Sequence: 1 MADKGMTYSFVRDNLVLR.....DMTPEIFDEDESEEDIDI 616

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_80.\*  
1: PIR1.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128	4.1	1556	2 S59393	probable membrane
2	125	4.0	1175	2 S70843	hemolytic protein
3	122	3.9	1203	2 B55094	chromosomal protei
4	121.5	3.9	1377	2 C70148	DNA-directed RNA p
5	119	3.8	547	2 AH0098	probable methyl-ac
6	118.5	3.8	1483	2 C97012	probably celluloso
7	117.5	3.8	1732	2 E71442	hypothetical prote
8	117.5	3.8	2253	2 T30336	nuclear/mitotic ap
9	116.5	3.8	1166	2 C96945	cation efflux syst
10	116.5	3.8	2269	2 T28677	rhoxy protein -
11	116	3.7	1163	2 P84669	probable chromosom
12	116	3.7	1166	2 T28680	fibrinogen-binding
13	115.5	3.7	1191	2 B97116	chromosome segrega
14	115.5	3.7	1198	2 B28065	Ca2+-transporting
15	115	3.7	1536	2 A43855	high-molecular-wei
16	114.5	3.7	833	2 AH2444	hypothetical prote
17	114	3.7	1640	2 D86798	prophage p13 prote
18	113.5	3.7	931	2 JC2542	alpha-2(B)-catenin
19	113.5	3.7	1819	2 D97033	uncharacterized pr
20	113.5	3.7	2052	2 C97038	phage-related prot
21	113	3.6	806	2 A41341	microbial serine p
22	112.5	3.6	4563	1 LPHUB	apolipoprotein B-1
23	112	3.6	621	2 A57591	Id-associated prot
24	111.5	3.6	584	2 E97303	ABC-type multidrug
25	111.5	3.6	1274	2 A89959	hypothetical prote
26	111	3.6	1141	2 B89824	hypothetical prote
27	110.5	3.6	1270	2 T51227	related to verruco
28	110.5	3.6	1290	2 S76853	hypothetical prote
29	110.5	3.6	1304	2 T19337	hypothetical prote

30	110	3.6	649	2 AF2866	methyl-accepting c
31	110	3.6	670	2 C97643	similar to mcp ge
32	109.5	3.5	683	2 D83511	flagellar hook-ase
33	109.5	3.5	1220	2 S13057	Ca2+-transporting
34	109.5	3.5	2570	2 T17451	fimbriae-associate
35	109.5	3.5	2660	2 B85822	probable invasin Z
36	109.5	3.5	3488	2 T34418	hypothetical prote
37	109	3.5	470	2 D69394	phosphoribosylamin
38	109	3.5	1810	2 B88481	protein C16A3.2 [I
39	109	3.5	1819	2 D97132	uncharacterized ph
40	109	3.5	1926	2 JC4842	DNA-binding nuclea
41	108.5	3.5	594	2 S37816	hypothetical prote
42	108.5	3.5	1081	2 T15692	hypothetical prote
43	108.5	3.5	1300	2 S73679	probable lipoprote
44	108.5	3.5	1335	2 G90975	probable factor [I
45	108	3.5	565	2 H83996	DNA repair and gen

## ALIGNMENTS

### RESULT 1

S59393  
Probable membrane protein YLR247c - yeast (Saccharomyces cerevisiae)  
N;Alternate names: hypothetical protein L9672.14  
C;Species: Saccharomyces cerevisiae  
C;Date: 30-Nov-1995 #sequence\_revision 23-Feb-1996 #text\_change 09-Jul-2004  
C;Accession: S59393  
R;Johnson, D.

submitted to the EMBL Data Library, February 1995  
A;Description: The sequence of S. cerevisiae cosmid 9672.  
A;Reference number: S59386

A;Accession: S59393

A;Molecule type: DNA

A;Residues: 1-1556 <JOH>

A;Cross-references: UNIPROT:Q06554; UNIPARC:UPI0000069EBE; EMBL:U20865; NID:G662330; PIR

A;Experimental source: strain S288C (AB972)

C;Genetics:

A;Gene: MIPS:YLR247C

A;Cross-references: SGD:S0004237

A;Map position: 12R

C;Keywords: transmembrane protein

F;1235-1282/Domain: RING finger homology <RRN>

F;1236-1252/Domain: transmembrane #status predicted <ITM>

Query Match 4.1%; Score 128; DB 2; Length 1556;  
Best Local Similarity 19.1%; Pred. No. 2.6;  
Matches 125; Conservative 111; Mismatches 194; Indels 224; Gaps 34;

QY	24	ATKSGIKISVREDRGTSLLQKAPAGTEDEFWEL----	DQVYVDKIRKPLEEEKMD 79
DB	778	ALRSPFKI--RDDLMTHIROK--FNVEDPDKSLNLSEDEHMDRFG--KETSSGDE 831	
QY	80	STR-VSGAVAAATERSVEPNFNSKEAANIEMAGVDDERAGGSLVDRNRKNG--VSN 135	
DB	832	SDEINGA-----KNHNNND-----GMLNHLKKGLRAMN 865	
QY	136	MAYNLSLFGVFPFALT-----PFSAILSEGEN-SIWQN----- 169	
DB	866	LHDCYFFLGSVVYNLGTRKLEADDKHKRKEKEEVVYSDVFPKNELEIEENLLQEN 925	
QY	170	--QQAIRIALADEDGKQTRTGGQRVDMA--DVTKLNVVTANGKQVENVL----ND 221	
DB	926	YANAETLRKSLSS-----ARKVDMTIKMTKFPAPMTSNIPRLINIEFHKND 976	
QY	222	LKAAPQSRPKRS-----DYRKGQSKA 244	
DB	977	YSSNLAVSRFCFSLSKLIEGLNEQTKNFNLLDELLIIIEPVHRTDDDSSTNKIIGNEE 1036	
QY	245	TESISNQ-----CMALIMK---SVLSADQLFAPGVKMMRTNGFNAGSVYTTIAEGANI 293	
DB	1037	YTSIDSQDKIFSLLCLEILLQNRDNLTSSES-----EVKIPK-----HLVPFGSI 1083	

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2006, 09:27:05 ; Search time 77 Seconds  
(without alignment)

5644.226 Million cell updates/sec

Title: US-10-734-782-4

Perfect score: 3097

Sequence: 1 MADKGMTYSPDVRNTLVVR.....DMTEPFEDEDEEEDIDI 616

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3097	100.0	616	Q91BR3_9ORTO	Q91br3 infectious
2	3083	99.5	616	Q8V3T7_9ORTO	Q8v3t7 infectious
3	3067	99.0	616	Q6YN41_9ORTO	Q6yn41 infectious
4	3029	97.8	616	Q88VU1_9ORTO	Q88vul infectious
5	152	4.9	1758	Q9JMS5_ECOLI	Q9jms5 escherichia
6	143.5	4.6	6084	Q87ID8_VIBPA	Q87id8 vibrio para
7	131.5	4.2	2089	Q82K38_RHIME	Q82k38 rhizobium m
8	130.5	4.2	794	Q7YX86_CRYPV	Q7yx86 cryptospori
9	129	4.2	1023	Q93T53_STRPY	Q93t53 streptococc
10	128.5	4.1	6077	Q91L85_WSSV	Q91l85 white spot
11	128	4.1	1556	Q06554_YEAST	Q06554 saccharomyc
12	127	4.1	566	Q74805_YERPE	Q74805 yersinia pe
13	127	4.1	566	Q8CZK7_YERPE	Q8czk7 yersinia pe
14	126	4.1	764	Q4HG87_CAMCO	Q4hg87 campylobact
15	126	4.1	957	Q600S0_MYCHY	Q600s0 mycoplasma
16	126	4.1	6077	Q8QTB7_WSSV	Q8qtb7 white spot
17	126	4.1	6077	Q8VAP1_WSSV	Q8vap1 white spot
18	125	4.0	1175	Q77955_HARDU	Q77955 haemophilus
19	125	4.0	1175	Q7VL76_HABDU	Q7vl76 haemophilus
20	124.5	4.0	1087	Q8W1X3_ARATH	Q8w1x3 arabidopsis
21	124.5	4.0	3165	Q8RD09_FUSNN	Q8rdg9 fusobacteri
22	123.5	4.0	892	Q6B2A1_DEBHA	Q6b2a1 debaryomyce
23	123	4.0	1314	Q7Q251_ANOGA	Q7q251 anopheles g
24	122	3.9	1203	1 SMC2_XENLA	P50533 xenopus lae
25	121.5	3.9	799	2 Q4PGW1_USTWA	Q4pgw1 ustilago ma
26	121.5	3.9	1377	1 RPOC_BORBU	Q5l349 borrelia bu
27	121.5	3.9	3620	2 Q5XW66_BARVI	Q5xw66 bartonella
28	121	3.9	862	2 Q5L5W4_CAEER	Q5l5w4 caenorhabdi
29	121	3.9	2402	2 Q3AER7_STAEP	Q3aer7 staphylococ
30	121	3.9	2402	2 Q5HKF4_STAEP	Q5hkf4 staphylococ
31	120.5	3.9	959	2 Q8S950_TOBAC	Q8s950 nicotiana t

32 120.5 3.9 1377 1 RPOC\_BORBA Q66in0 borrelia ga  
33 120 3.9 626 2 Q46737\_ECOLI Q46737 escherichia  
34 120 3.9 703 2 Q4H8L2\_9DEIO Q4h8l2 deinococcus  
35 120 3.9 1083 2 Q7PMF8\_ANOGA Q7pmf8 anopheles g  
36 119.5 3.9 727 2 Q64JU8\_PLAVI Q64ju8 plasmodium  
37 119.5 3.9 727 2 Q64JU3\_PLAVI Q64ju3 plasmodium  
38 119.5 3.9 3187 2 Q4WQ48\_ASFFU Q4wq48 aspergillus  
39 119 3.8 547 2 Q8ZHT2\_YERPE Q8zht2 yersinia pe  
40 119 3.8 547 2 Q667E2\_YERPS Q667e2 yersinia ps  
41 119 3.8 727 2 Q64JU7\_PLAVI Q64ju7 plasmodium  
42 119 3.8 872 2 Q5ACU0\_CANAL Q5acu0 candida alb  
43 118.5 3.8 1042 2 Q419R7\_GIBZS Q419r7 gibberella  
44 118.5 3.8 1090 2 Q91E95\_9RBOV Q91e95 human rotav  
45 118.5 3.8 1093 2 Q5XAS7\_STRP6 Q5xas7 streptococc

#### ALIGNMENTS

##### RESULT 1

ID Q91BR3\_9ORTO PRELIMINARY; PRT; 616 AA.  
AC Q91BR3;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Putative nucleoprotein.  
OS Infectious salmon anemia virus.  
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae; Isavirus.  
OX NCBI\_TaxID=55987;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=21342912; PubMed=11450947; DOI=10.1023/A:101110105919;  
RA Ritchie R.J., Heppell J., Cook M.B., Jones S., Griffiths S.G.;  
RT "Identification and characterization of segments 3 and 4 of the ISAV genome.";  
RL Virus Genes 22:289-297 (2001).  
DR EMBL; AF306549; AAK97306.1; -, mRNA.  
DR InterPro; IPR006089; Acyl-CoA dh.  
KW PROSITE; PS00073; ACYL\_COA\_DH\_2; UNKNOWN\_1.  
KW Viral nucleoprotein.  
SQ SEQUENCE 616 AA; 68050 MW; 2436D70119524E74 CRC64;

Query Match 100.0%; Score 3097; DB 2; Length 616;

Best Local Similarity 100.0%; Pred. No. 9.3e-195;

Matches 616; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MADKGMTYSPDVRNTLVVRSTATKSGIKISYREDRGTSLLQKAPAGTDEFWVELDQD 60  
Db 1 MADKGMTYSPDVRNTLVVRSTATKSGIKISYREDRGTSLLQKAPAGTDEFWVELDQD 60  
Qy 61 VYVDKKIRKFLERKKMDSTRVSGAVAAAIERSVEFDNFSKEAAANIENAGVDDERAGG 120  
Db 61 VYVDKKIRKFLERKKMDSTRVSGAVAAAIERSVEFDNFSKEAAANIENAGVDDERAGG 120  
Qy 121 SGLVDRNRKNGVSNMAYNLSLFIQVFPALTTFSSAILSEGENSIWNGQAIIRILALA 180  
Db 121 SGLVDRNRKNGVSNMAYNLSLFIQVFPALTTFSSAILSEGENSIWNGQAIIRILALA 180  
Qy 181 DEPGKQRTFGGQVRDMADVTKLVNVTYANGKVKQVEVNLNDLKAAPRQSRPKRSDYRKQ 240  
Db 181 DEPGKQRTFGGQVRDMADVTKLVNVTYANGKVKQVEVNLNDLKAAPRQSRPKRSDYRKQ 240  
Qy 241 GSATATSSISNQCALIMKSVLSADQLFAPGVQMTGTNGFNASVYTLAEGANIPSKYLRLH 300  
Db 241 GSATATSSISNQCALIMKSVLSADQLFAPGVQMTGTNGFNASVYTLAEGANIPSKYLRLH 300  
Qy 301 MRNCGGVALLDMGKKIKNSPEGAKSKIPSIIOKKVGRGCRTEFORLLTSALKISDGENK 360  
Db 301 MRNCGGVALLDMGKKIKNSPEGAKSKIPSIIOKKVGRGCRTEFORLLTSALKISDGENK 360  
Qy 361 FORIMDTLCTSFIDPPRTTKCFIPPTSSIMTYIQEGNSVLMDFMKNGDACKICREAK 420

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

M protein - protein search, using sw model

un on: January 13, 2006, 09:28:46 ; Search time 9 Seconds  
(without alignments)  
647.088 Million cell updates/sec

File: US-10-734-782-4  
erfect score: 3097  
Sequence: 1 MADKGMTYSFVDRNTLVVR.....DMPRIEDEDDEDEDIDI 616

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

searched: 67062 seqs, 9454214 residues

total number of hits satisfying chosen parameters: 67062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA New:
- 1: /cgm2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB\_PEP.\*
  - 2: /cgm2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB\_PEP.\*
  - 3: /cgm2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB\_PEP.\*
  - 4: /cgm2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB\_PEP.\*
  - 5: /cgm2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB\_PEP.\*
  - 6: /cgm2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB\_PEP.\*
  - 7: /cgm2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB\_PEP.\*
  - 8: /cgm2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB\_PEP.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	3997	100.0	616	7	US-11-083-800-4
2	116.5	3.8	1086	7	US-11-052-554A-380
3	113.5	3.7	2053	7	US-11-013-759-9
4	112.5	3.6	1572	6	US-10-793-626-2906
5	111	3.6	1992	7	US-11-013-759-3
6	111	3.6	1992	7	US-11-013-759-13
7	111	3.6	2047	7	US-11-013-759-4
8	111	3.6	2047	7	US-11-013-759-7
9	108.5	3.5	1300	7	US-11-052-554A-125
10	105	3.4	2314	7	US-11-013-759-11
11	103	3.3	5291	7	US-11-052-554A-281
12	102	3.3	1976	7	US-11-069-834-52
13	102	3.3	1976	7	US-11-069-834-54
14	101	3.3	446	7	US-11-196-475-154
15	101	3.3	531	7	US-11-060-914-4
16	101	3.3	2233	6	US-10-873-528-2
17	100.5	3.2	522	6	US-10-793-626-456
18	100.5	3.2	522	6	US-10-821-234-1188
19	100	3.2	930	6	US-11-196-475-152
20	99.5	3.2	453	6	US-11-196-475-1641
21	99.5	3.2	1970	6	US-10-821-234-1641
22	99.5	3.2	2902	7	US-11-052-554A-91
23	99.5	3.2	3256	7	US-11-124-368A-304
24	98	3.2	446	7	US-11-196-475-150
25	97.5	3.1	424	6	US-10-485-517-264

26	97.5	3.1	424	6	US-10-485-517-281
27	97.5	3.1	894	6	US-10-485-517-416
28	97	3.1	1960	7	US-11-069-834-50
29	96.5	3.1	453	7	US-11-196-475-146
30	96	3.1	588	7	US-11-196-475-122
31	96	3.1	794	6	US-10-793-626-1050
32	95	3.1	943	6	US-10-467-657-5508
33	94.5	3.1	446	7	US-11-196-475-158
34	94.5	3.1	1327	7	US-11-052-554A-13
35	94	3.0	585	6	US-10-510-386-20
36	94	3.0	594	6	US-10-510-386-38
37	94	3.0	1571	7	US-11-052-554A-2
38	94	3.0	1960	7	US-11-069-834-48
39	93	3.0	450	7	US-11-196-475-148
40	93	3.0	453	7	US-11-196-475-156
41	92.5	3.0	573	7	US-11-196-475-112
42	92.5	3.0	1411	6	US-10-453-372-1106
43	92	3.0	757	7	US-11-052-554A-378
44	92	3.0	2896	7	US-11-124-368A-306
45	91.5	3.0	1706	7	US-11-052-554A-176

## ALIGNMENTS

### RESULT 1

US-11-083-800-4  
Sequence 4, Application US/11083800  
Publication No. US20050261227A1  
GENERAL INFORMATION:  
APPLICANT: Griffiths, Steven  
APPLICANT: Ritchie, Rachael Jane  
TITLE OF INVENTION: Nucleic Acid And Amino Acid Sequences Of Infectious Salmon Anaemia Virus And Their Uses As Vaccines  
TITLE OF INVENTION: Infectious Salmon Anaemia Virus And Their Uses As Vaccines  
FILE REFERENCE: H-32318D  
CURRENT APPLICATION NUMBER: US/11/083,800  
PRIOR FILING DATE: 2005-03-18  
PRIOR APPLICATION NUMBER: US 10/049,086  
PRIOR FILING DATE: 2002-02-06  
PRIOR APPLICATION NUMBER: PCT/GB00/02976  
PRIOR FILING DATE: 2000-08-07  
PRIOR APPLICATION NUMBER: GB 0006674.6  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: GB 0005848.7  
PRIOR FILING DATE: 2000-03-11  
PRIOR APPLICATION NUMBER: GB 9918588.6  
PRIOR FILING DATE: 1999-08-07  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 616  
TYPE: PRT  
ORGANISM: Infectious Salmon Anaemia Virus  
US-11-083-800-4

Query Match 100.0%; Score 3097; DB 7; Length 616;

Best Local Similarity 100.0%; Pred. No. 3.8e-248;

Matches 616; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MADKGMTYSFVDRNTLVVRRTATKSGIKISYREDRGTSLLQKAFAGTDEFWVLDQD	60
DB	1	MADKGMTYSFVDRNTLVVRRTATKSGIKISYREDRGTSLLQKAFAGTDEFWVLDQD	60
QY	61	VYVDKIRKFLBEEKKMDSTRVSGAVAAIERSVEFDNFSKEAAANIENAGVDDREAGG	120
DB	61	VYVDKIRKFLBEEKKMDSTRVSGAVAAIERSVEFDNFSKEAAANIENAGVDDREAGG	120
QY	121	SGLVDRNRKNGVSNMAYNLSLFIQVFPALTFFSAILSEGEMSIWNGQAIRILALA	180
DB	121	SGLVDRNRKNGVSNMAYNLSLFIQVFPALTFFSAILSEGEMSIWNGQAIRILALA	180
QY	181	DEGDKRTRTGGORVDMADYTKLVNVTANKVKQVFNLDLKAAFQSPKPSDYRKGQ	240